



Implementation of Electronic Transmission of Chemical Occurrence Data (CFP/EFSA/DATEX/2009/01) in Denmark

Andersen, Jens Hinge

Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Andersen, J. H. (2011). *Implementation of Electronic Transmission of Chemical Occurrence Data (CFP/EFSA/DATEX/2009/01) in Denmark*. European Food Safety Authority.
<http://www.efsa.europa.eu/en/supporting/pub/152e.htm>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

SCIENTIFIC/TECHNICAL REPORT submitted to EFSA

Implementation of Electronic Transmission of Chemical Occurrence Data (CFP/EFSA/DATEX/2009/01) in Denmark¹

Prepared by Jens Hinge Andersen, National Food Institute, DTU

Abstract

The present report describes the resources used and challenges meet in restructuring the Danish national data repositories for chemical contaminants and pesticides and in coding these according to the EFSA “Standard Model”.

Summary

EFSA has prepared a standard data model for the transmission of chemical occurrence data and pesticide residues. This model is referred to as the “Standard Model” (SM) or the “Standard Sample Description” (SSD). The aim of the present project was to collect and transform data from the Danish national data repositories for chemical contaminants and pesticides and in coding these according to the SM and to facilitate the continued use of the SM in answering future EFSA data calls in these areas. Restructuring was performed using existing equipment and programming tools. SAS procedures and Excel mapping tables were developed to handle the data transformation. The main challenges were mapping existing data to the SM, synthesising necessary information missing in the national repository, and dealing with the complexity of pesticide substance reporting.

Key words: Chemical, Occurrence, Residues, Transmission, XML, Data, Pesticides, Grant, DTU

¹ Question No EFSA-Q-2009-00838. Accepted for Publication on 25 April 2011

Table of Contents

Abstract	1
Summary	1
Table of Contents	2
Background	3
Acknowledgements	3
Introduction and Objectives	4
1. Restructuring the national data repositories	4
1.1. Existing systems	4
1.1.1. Danish Veterinary and Food Administration (DVFA).....	4
1.1.2. National Food Institute (DTU Food).....	5
1.2. Project implementation	5
1.3. Future system.....	5
2. Overview of the mapping strategy.....	6
3. Populating the SSD.....	6
4. Maintenance	7
4.1. Mapping tables.....	7
4.2. SAS code.....	7
5. Challenges	7
5.1. Maintenance of dynamic mapping tables	7
5.1.1. Product code and related fields	7
5.1.2. Parameter code and related fields.....	8
5.1.3. Sampling point.....	8
5.2. Area specific demands	8
5.2.1. Pesticides.....	8
5.2.2. PCBs and Chlorinated persistent pesticides	9
5.2.3. Mycotoxins	9
5.2.4. Heavy metals	9
5.2.5. PAHs.....	10
5.2.6. Dioxins	10
5.2.7. Acrylamide	10
6. Limitations and enhancement to the SSD	10
7. Cost/benefit analysis of data calls	11
8. Improving national repository towards SSD compliance.....	11
Results	13
Conclusions and Recommendations.....	14
Appendices	15

Background

Data collection is an important task of EFSA² and a fundamental component of risk assessment (Articles 22 and 23 of Regulation EC No 178/2002). As a result, EFSA receives an increasing volume of data from different data providers in support of its scientific activities. This is particularly true for data collections of concentrations of chemical contaminants and pesticide residues from European and national control monitoring programs that generate large volumes of data. Such large volumes are difficult to manage without a standardised and structured approach.

In order to address this challenge EFSA has prepared a standard data model for the transmission of chemical occurrence data based on proposals received within the Article 36 grant “Development of a Standard Food Classification and Sample Description System for Chemical Occurrence Data Storage” (CFP/EFSA/DATEX/2007/02). The proposed standard data model has been enhanced in order to also support the transmission of pesticide residues and to be compatible with the EFSA IT systems. The draft standard model will be referred to as the “Standard Model” (SM) or the “Standard Sample Description” (SSD).

Acknowledgements

This contract/grant was awarded by EFSA to:

Danish Technical University

National Food Institute

Dept. of Food Chemistry

Mørkhøj Bygade 19

2860 Søborg

Denmark

Contract/grant title: Electronic Transmission of Chemical Occurrence Data

Contract/grant number: CFP/EFSA/DATEX/2009/01

² REGULATION (EC) No 178/2002 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, of 28 January 2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety OJ L 31, 1.2.2002, p. 1–24.

Introduction and Objectives

The project Electronic Transmission of Chemical Occurrence Data (CFP/EFSA/DATEX/2009/01) has been financially supported by EFSA according to Article 36 of the European Parliament and Council Regulation (EC) No 178/2002.

The outcome of the project will assist EFSA to collect, collate, analyse and summarise relevant scientific and technical data in the fields within its mission and have been achieved in close collaboration with EFSA DATEX and PRAPeR units.

During the project, selected data from the Danish Veterinary and Food Administration (DVFA) from national monitoring of chemical occurrences and pesticide residues have been made available for transmission to EFSA in accordance with EFSA's Standard Model (aka EFSA Standard Sample Description, SSD). Substance groups included so far are: Pesticides, PCBs and chlorinated persistent pesticides, mycotoxins, heavy metals, PAHs, dioxins, acrylamide and 3-MCPD.

Data have been collected by DVFA regional laboratories and validated and transformed by the National Food Institute (DTU Food).

During the project test phase, data from three selected areas (PAH, heavy metals and mycotoxins) have been transmitted successfully using the EFSA Data Collection Framework (DCF). In addition, data calls for acrylamide and pesticides have been transmitted using DCF.

1. Restructuring the national data repositories

1.1. Existing systems

1.1.1. Danish Veterinary and Food Administration (DVFA)

Primary data from samples and results are stored at the Danish Veterinary and Food Administration (DVFA) in a LIMS system (Oracle based). The system is a combination of an administrative journal system, ScanJour, and a laboratory information system, SQL*LIMS. The sample information and analytical results are stored in the two regional laboratories of the Food Administration. The LIMS systems of these laboratories are clones of the same system.

Entry of sample information is performed by sampling officers through the administrative journal system that also contains information on inspections and follow-up measures. Sample information is transferred electronically to the laboratory information system of the responsible DVFA laboratory.

Entry of method and result data is performed by the laboratory. A report (pdf-format) on the analytical results is sent back to the sample file in the journal system after approval of results.

When samples are analysed by external laboratories, one DVFA laboratory is responsible for result entry into the LIMS.

Both systems are heavily modified standard systems and it is generally agreed that further modifications – especially of the LIMS – is not recommended.

1.1.2. National Food Institute (DTU Food)

SAS software and the Windows Office Suite constitute the basis for managing information from laboratory data and for statistical analysis of scientific data at DTU Food. One example is collecting and reporting data from the national pesticide residue monitoring programs using applications build and maintained by DTU Food personnel.

1.2. Project implementation

Due to the inflexibility of the existing LIMS, it was decided to build an interface system to capture data from the LIMS and transform these to the EFSA Standard Sample Description (SSD³).

Since the data structure in the LIMS and the SSD is not directly comparable, it was decided that the interface should be batch wise in order to cope with unforeseen complications. This choice was also based on the fact that data from different chemical domains (at least chemical contaminants and pesticides) was to be managed by the same interface, so flexibility had a very high priority.

The interface was build using existing programming tools and computers, thus the resources needed were mainly human.

The fundamental parts of the interface programming were implemented through the EFSA pilot project on the reporting of pesticide residues data 2008. Some changes to the SSD (including handling of validation rules) was introduced with the 2010 version of the Standard Sample Description, but a large part of the resources needed was for building mapping tables and analysing data in order to comply with the differences between data demands from different domains of chemical contaminants and pesticides.

1.3. Future system

Since the onset of the present project, DVFA has initiated planning of a new LIMS system to replace the existing system. Experience gained from the present project will be available for the new project already from the design phase.

Unfortunately, clear specifications for the future EFSA Food Classification system is not yet available. In addition, the SSD may not yet have reached a stable form, since only few data collections have been performed. One important aspect from the point of data providers is to which degree it will be possible to minimise differences in data specifications between the different domains using the SSD.

³ Standard Sample Description, <http://www.efsa.europa.eu/en/supporting/pub/1457.htm>

2. Overview of the mapping strategy

Data is extracted from the LIMS by the responsible laboratory using Crystal Reports to generate Excel-files with the same structure as the SSD, e.g. each line describing a result with associated information on result, sample, project etc. (referred to below as the 'LIMS-XL extract').

Data collection is performed batch wise. Selection of extracted results is based on project ID. Results for the extracts are defined as those results that were issued on approved analysis certificates from the specified project.

The extraction programme has been developed by the DVFA laboratory in Region East based on specifications from DTU Food during the EFSA Pilot Project on the reporting of pesticide residues data 2008. During the implementation of the present project, it have been modified and tested for the reporting of pesticide residues 2009 according to the 2010 version of the Standard Sample Description, and a few more fields from the LIMS have been added to the extracted data structure in order to cope with some of the areas of chemical contaminants.

Excel tools have been developed to assist the laboratory in validating and correcting data, maintaining traceability of such changes. Correction of data in the LIMS system at this stage has been found counterproductive.

A suite of translation tables are maintained in Excel and imported to SAS. These tables also include information not supplied by the LIMS.

Information from these mapping tables is joined with and/or replaces information from the LIMS extract.

3. Populating the SSD

Most fields in the SSD are populated by joining through mapping tables on relevant level or by calculation from fields in the LIMS extract.

This mapping is detailed in appendices A and B.

During the process of including different data domains in the SSD it was found useful to include a few extra SSD columns for manual entry in the 'LIMS-XL extract':

- prodCom (product comment): PAHs, acrylamide
- resLOD (LOD): Heavy metals
- origFishAreaText (fishing area): Heavy metals, PCB and OCPs, dioxin
- sampPoint (sampling point): Acrylamide, pesticides
- progSampStrategy (sampling strategy): Mycotoxins.

This information is not stored in the LIMS in a way suited for automatic extraction. Sampling point and sampling strategy can in most cases be associated with project ID in an Excel mapping table.

4. Maintenance

4.1. Mapping tables

Mapping tables are maintained in Excel 2007/2010 (stored as Excel 97-2003 files due to constraints in the installed SAS-version). Values from the Danish LIMS are translated to EFSA codes. For verification of codes, textual translation of EFSA codes is visible in the Excel file through lookups in the relevant EFSA catalogue.

4.2. SAS code

Most code related to the SSD structure is maintained in Excel and copied to SAS macros. The current Excel version of the SSD (Variables sheet) is directly used as the defining basis for building generic formulas to generate field specific SAS code.

This structure has so far proven both flexible and safe for implementing changes in SSD, in validation rules and in mapping politics.

5. Challenges

Those parts of the SSD that have direct counterparts in the LIMS system, and therefore can be mapped by using stable mapping tables have been rather unproblematic to implement and maintain.

The major challenges encountered can be broken down to two categories: Maintenance of dynamic mapping tables (i.e. tables with information that needs to be updated for most individual data calls) and area specific demands (i.e. special mapping of information not contained in the LIMS or placed as free text in different LIMS data fields).

5.1. Maintenance of dynamic mapping tables

5.1.1. Product code and related fields

The product catalogue of DVFA contains a large collection of products not relevant for data collection of chemical contaminants or pesticides. For this reason it was decided to create the mapping table dynamically as needed. The demand for new product codes is of course greatest the first time a new data domain is included. But so far, it has been necessary to initiate an update process for each data call or domain.

In general, coding for products in pesticide projects has been rather unproblematic, partly because these products mainly have been unprocessed products of single agricultural commodities, partly because a specialized catalogue (MATRIX) has been available.

For other products, a few issues have been problematic:

- The granularity of the two systems are very different (e.g. LIMS have “nudler” (“noodles”); FOODEX has “Glass noodle”, “Noodle, rice”, “Noodle, wheat flour, with eggs” and “Noodle, wheat flour, without eggs” – so where to put “nudler”?)
- Lack of knowledge of English food terms – and missing explanations in the food catalogues (i.e. what is “Glass noodle” made from? How is “Calabrese salami” different from “Italian-type salami”? (and where to put Danish produced “Salami”⁴?)
- In some cases, the sample may be less precisely described in LIMS catalogue categories, but detailed in free text fields (i.e. in the acrylamide project samples of partly processed potatoes were collected. Additional information could be added in free text (i.e. “half-baked and frozen, to be baked or fried by consumer” (in Danish)). Although such information could be transferred by hand to the field prodText or prodCom, this cannot be expected except in singular cases.

5.1.2. *Parameter code and related fields*

The coding of parameters/analytes has been rather unproblematic except for pesticides, where coding have been complex and time consuming. This has been described in further details below.

5.1.3. *Sampling point*

Information on sampling point is stored in the register of inspection points in the journal system of DVFA, and would need special programming for inclusion in the ‘LIMS-XL extract’. Due to lack of resources at the DVFA for this work, information on sample point had to be linked manually to each project in an Excel mapping file. Where this information was not satisfactory, information in an additional column in the ‘LIMS-XL extract’ had to be provided manually.

5.2. *Area specific demands*

5.2.1. *Pesticides*

A major challenge in the pesticide domain has been the requirement to report all single results, including compounds included in the analysis, but not detected in the sample. In the LIMS, only the method used and the residues identified are recorded with the sample. Thus, the individual non-detects have to be synthesized. This is done by SAS procedures using method information in an Excel mapping file.

Another challenge has been the complex residue definitions where results must be reported as the sum of several individual residues analysed for in the laboratory. In addition, some of the residue definitions (i.e. which individual residues to sum) vary between different sample types.

⁴ Except as FOODEX-code ~ ”Sausage” and prodText = “Salami”

As such, this is not an problem originating from the SSD, and to some extent, the strict requirement from the SSD have improved the quality of the data at the expense of increased investment in validation of data from the laboratory.

Although information on substances and LOQs is stored in the LIMS, it has not been possible to access this information automatically, and information on methods have to be handle through the Excel sheet that also maps the information to SSD on parameter codes etc.

In the domain of pesticides, methodology has been relatively less stable and more complex than in many other domains: New substances are added during the year, or are moved from e.g. GC-methods to LC-methods. These circumstances add to the workload in transforming LIMS data to SSD data.

5.2.2. *PCBs and Chlorinated persistent pesticides*

For persistent organic chlorine pesticides (OCPs) only methods and detected compounds are recorded, so results for not detected compounds must be synthesized as for the other pesticides. However, there have been fewer changes in methods and compounds than for the other pesticides.

Marine samples require the recording of the place where the fish was caught. This information is stored in free text in the LIMS, but for different projects and data domains both the field and the format may vary. The information have been stored as geographical names of the waters (typical for inland or local sea waters), as FAO or ICES IDs and/or as geographical coordinates. In order to catch this information, a new column in the 'LIMS-XL extract' has been dedicated to store text for the SSD variable origFishAreaText. The 'LIMS-XL extract' has then been searched manually for relevant information (in LIMS columns 'OprindelsesstedBeskrivelse', 'Udtagningsstedbeskrivelse', 'PROJEKTOPLYSNING1...4'). The information has been transferred as it is without translation.

This procedure has been used for marine samples in other data domains as well.

5.2.3. *Mycotoxins*

Aflatoxin monitoring requires that subsamples are taken and recorded in the LIMS. Unfortunately, handling of subsamples is not implemented in the LIMS, where every subsample is identified as an independent sample. However is was possible to add a column in the 'LIMS-XL extract' from Crystal Reports with information on external sample labelling making a transformation of subsamples into samples with subsample IDs possible.

5.2.4. *Heavy metals*

For heavy metals, information on LOQ and LOD from the period 2004-2009 was dependent on project ID and metal. This information was not available in the 'LIMS-XL extract' but was available on a separate, manually created, Excel sheet. LOQ was linked (on sample ID, project ID and compound) to the column 'LGM' in the 'LIMS-XL extract' and an extra column were created for LOD.

5.2.5. *PAHs*

In this domain, detailed knowledge on product treatment can be essential in order to evaluate analytical results. This information could be stored as free text in the LIMS, but was instead recorded separately in Excel sheets together with sample ID. Since the format of the Excel sheets have been essentially constant through the period from which results have been collected so far (2006-09), it has been possible to add the information to SSD by adding a column for the free text field prodCom (product comment) in the 'LIMS-XL extract' and linking on sample ID (project by project) to the Excel sheets with information on treatment.

5.2.6. *Dioxins*

Results below LOQ are not included on the analysis certificate for dioxins. Therefore the standard 'LIMS-XL extract' has to be supplemented with data from another Crystal Reports extract. These data are then linked to the standard 'LIMS-XL extract'.

5.2.7. *Acrylamide*

As mentioned previously, sample types and product treatment was difficult to match to the FOODEX and PRODTR catalogues.

6. **Limitations and enhancement to the SSD**

Reaching a balance between coded fields and free text fields is problematic. Only coded fields can be expected to be mapped automatically from LIMS to SSD.

The existing Food Classification system is clearly not adequate as a general food catalogue. Hopefully its successor that is under way from the EFSA working group will provide better coverage and guidance to the mapping process.

The catalogue for product processing may be integrated in the coming food classification catalogue. At its present state the catalogue is weak e.g. in specifying processes for dairy and pre-processed products.

Data calls are issued from different branches/data domains of EFSA with different needs for sample description and result information. On the other hand, a laboratory system very often have to deal with samples and results from all fields relevant for the SSD. It would not be acceptable if different mappings were specified for different areas. In order to integrate the requirements from SSD into a LIMS, it is of outmost importance, that data requirements are compatible between the different SSD domains and that a common understanding of the SSD can be maintained.

Valuable efforts have been put into constructing the SSD and in implementing reporting interfaces at data providers. It cannot be expected that demands and means for this system be unchanged if the system should have a lifespan beyond a few years. For this reason proper resources must be allocated – especially at EFSA – to maintain the system and preserve knowledge through these changes.

7. Cost/benefit analysis of data calls

A thorough cost/benefit analysis on the efforts put into answering EFSA data calls is difficult from the present experience as repetitive calls so far only have been answered for the pesticide domain.

Answering the first call within a data domain using the SSD is certainly time-consuming, as new entries to catalogues have to be created and sometimes special recording needs must be addressed. It can be expected that additional calls will require less investment in catalogue updating.

In the pesticide domain, two calls have been answered (and a third is under preparation). Pesticide data have for many years been collected on a yearly basis for national reporting and for reporting agglomerated data to the EU Commission in a fixed format, and data validation have been a major part of that process.

The SSD model has a high degree of stringent data structure and validation rules, which have increased the workload during the process of getting data ready for reporting.

On the other hand, the external pressure for detailed formalised data in a presumably stable environment has increased the possibility to specify and push forward the manual validation routines in the process. The reporting laboratory (Region East) and the reporting organisation (DTU Food) has in cooperation specified more precise requirements for validation and developed tools for this purpose. Already this seems to have improved the quality of the sample and result data. It is expected that this process also will have a positive impact on the quality of method and substance catalogue data.

In addition, manual resources were needed in order to transform the validated national data to the older formats for reporting to the EU Commission. In the SSD model, the validated data are transformed to the required XML format automatically by SAS procedures.

So, from the point of view of the reporting organisation, the implementation of the SSD model is expected to be highly beneficial in the pesticide domain.

For other data domains the situation seems less clear. One important drawback in the Danish solution is the need for manual updating of catalogues. While most catalogues can be maintained by a central unit in the reporting organisation, maintenance of the substance and project catalogues require intimate knowledge of the specific domain. Since such updating typically is handled only once a year, adequate information on the procedure must be communicated.

On the other hand the collection of important national residue and contaminant data in a uniform detailed data format could have beneficial impact on national reporting and research into the collected data.

8. Improving national repository towards SSD compliance

As mentioned before, changes to the existing DVFA LIMS cannot be expected, but planning for a new LIMS have been initiated.

In order to facilitate transformation of LIMS data to the SSD model, some points should be considered for the design of a new LIMS, as this information has been missing in an adequate form in the existing LIMS:

- Additional project related information for the SSD should be included⁵.
 - progLegalRef (Programme legal reference)
 - progSampStrategy (Sampling strategy)
 - progType (Type of sampling program)
 - sampMethod (Sampling method)
- Additional sample related information for the SSD should be included.
 - origFishAreaCode (Area of origin for fisheries or aquaculture activities code)⁶
 - origFishAreaText (Area of origin for fisheries or aquaculture activities text)
 - procCountry (Country of processing)
 - sampPoint (Sampling point)
 - Codification of several existing free text fields should be considered.
- Additional laboratory related information from the SSD should be included.
 - labAccred (Laboratory accreditation)
 - labCountry (Laboratory country)
- Additional result related information from the SSD should be included.
 - LOQ, LOD, CC α , CC β
 - Recovery (where results have been corrected for recovery by calculation)
- Handling of subsamples should be enabled.
- Handling of information on Fat% and Moist% on sample level and expression of result (whether result is on whole weight, fat or dry matter).
- Handling of complex residue definitions.
- Information on follow-up actions on non-compliant results⁷ should be accessible for SSD extracts.
- Method data should include SSD relevant information and should be accessible for SSD extract.
- As many mapping catalogues as possible should be implemented.

⁵ In some projects, this information might be related to individual samples, not to all samples in the project.

⁶ The catalogue for fishing area (FAO Fisheries areas) in general lacks useful details for Danish samples.

⁷ In some domains (e.g. pesticides) information on follow-up actions are required also for some compliant results (not just results significantly above the legal limit but also for results numerically above the legal limit).

Most samples from DVFA are selected as an integral part of food inspectors' obligation to evaluate food safety at the individual establishment/sampling point, and sampling officers are trained and familiar with the requirements for such sampling. Sample information in relation to the establishment and to sample compliance has high priority.

In contrast to this, most samples from projects that generate data to the EFSA SSD has a monitoring aim being used for evaluating food safety for the general population. In this case, sample description should be as precise and detailed as prescribed for the project.

These two different points of focus should both be met by the DVFA LIMS, and tools should be available for sampling officers to select samples and record sampling information in accordance to the specific monitoring purpose of the project. One tool could be sample information entry forms that could be modified according to prescriptions of the project and project defined pick lists.

At present, during the validation process, minor errors are found in the data, especially in sample information. Most often these errors have little or no implications for evaluating food safety in relation to the sampling point, but may be important for monitoring statistics. Errors of this type sometimes hide when individual samples are reviewed, but become obvious when information from many samples from a project is aligned. Tools for handling this kind of error correction in a traceable manner should be available.

Results

During the project, selected data from the Danish Veterinary and Food Administration (DVFA) from national monitoring of chemical occurrences and pesticide residues have been made available for transmission to EFSA in accordance with EFSA's Standard Model (aka EFSA Standard Sample Description, SSD). Substance groups included are: Pesticides, PCBs and chlorinated persistent pesticides, mycotoxins, heavy metals, PAHs, dioxins, acrylamide and 3-MCPD.

Data have been collected by DVFA regional laboratories and validated and transformed by the National Food Institute (DTU Food).

During the project test phase, data from three selected areas (PAH, heavy metals and mycotoxins) have been transmitted successfully using the EFSA Data Collection Framework (DCF). In addition, data calls for acrylamide and pesticides have been transmitted using DCF.

Conclusions and Recommendations

CONCLUSIONS

A system for collecting and transforming data from the Danish Veterinary and Food Administration (DVFA) from national monitoring of chemical occurrences and pesticide residues have been implemented in order to make these data available for transmission to EFSA in accordance with EFSA's Standard Model in the prescribed XML format.

The system has been constructed as an interface layer between the DVFA LIMS as the existing LIMS cannot be changed.

RECOMMENDATIONS

All EFSA units requesting data in the Standard Model format should strive to maintain as uniform data requirements as possible in order to allow a common handling of transformation by the data providers who may have to provide data to different EFSA units or data domains.

If a new DVFA LIMS will be available in the future, much of the functionality for collecting and transforming data should be integrated in such new system. Experience gained through the present project should be used in the design of such integration.

Appendices

APPENDIX A: LIMS-XL EXTRACT

SAS dataset: LIMS	
PROJEKTNR	Unique project code (for all DK projects)
PROVENR	Unique sample code (for all DK samples)
Eksternt prøvenummer	External sample number (ID from sampling institution)
AnalyseLab	Analysing laboratory
UdtDato	Sampling date (DD-MM-YYYY)
FAERDIGDATO	Analysis date (completion) (DD-MM-YYYY)
Oprindelsesland	Country of origin
OPRINDELSESLANDKode	Code for country of origin
OprindelsesstedBeskrivelse	Free text description of origin
Udtagningsstedbeskrivelse	Free text description of sampling place
VAREARTDK	Food description from food list
HANDELSNAVN	Trade name (free text)
BEHANDLING1DK	Facet (method of conservation)
BEHANDLING2DK	Facet (part of item sampled)
BEHANDLING3DK	Facet (degree of comminution)
SPECIELORDNING1	Information of special legislation relevant product type (e.g. organic; GMO) is stored here.
SPECIELORDNING2	Information of special legislation relevant product type (e.g. organic; GMO) is stored here.
SPECIELORDNING3	Information of special legislation relevant product type (e.g. organic; GMO) is stored here.
PROJEKTOPLYSNING1	Free text fields. Only used if recording of project-specific information has been specified in project plan
PROJEKTOPLYSNING2	Free text fields. Only used if recording of project-specific information has been specified in project plan
PROJEKTOPLYSNING3	Free text fields. Only used if recording of project-specific information has been specified in project plan
PROJEKTOPLYSNING4	Free text fields. Only used if recording of project-specific information has been specified in project plan
DatoHoldbarhed	Expiry date (DD-MM-YYYY)
DatoProduktion	Production date (DD-MM-YYYY)
COMPONENT	Analytical substance/property/organism
Akkrediteret	Result by accredited method? (Y/N)
SD	Result standard deviation
TEXT_VALUE	Result (as text)
NUMBER_VALUE	Result (as number)
UNITS	Unit (result, LOQ etc.)
METHOD	Code and name of method
MRL	Legal limit
LGM	Lowest limit of measurement
BemTxt1	Remark to Konkl
Konkl	Remark #1 (code) to result

KonkTxt1	Remark #1 (translation of code) to result
BemTxt2	Remark to Konkl2
Konk2	Remark #2 (code) to result
KonkTxt2	Remark #2 (translation of code) to result

APPENDIX B: POPULATING THE SSD**Result level**

SSD-ID	Source	Mechanism	SAS dataset: EFSA
S.01	PROVENR		labSampCode
S.06	OPRINDELSESSESLANDKode	Mapping table EFSAland	origCountry
S.12	VAREARTDK, BEHANDLING1DK, BEHANDLING2DK, BEHANDLING3DK	Mapping table EFSAvare	EFSAProdCode
S.13	VAREARTDK, BEHANDLING1DK, BEHANDLING2DK, BEHANDLING3DK	Mapping table EFSAvare	prodCode
S.14	VAREARTDK, BEHANDLING1DK, BEHANDLING2DK, BEHANDLING3DK	Mapping table EFSAvare	prodText
S.15	OKO (see note)	Mapping table EFSAprodmd	prodProdMeth
S.17	VAREARTDK, BEHANDLING1DK, BEHANDLING2DK, BEHANDLING3DK	Mapping table EFSAvare	prodTreat
S.22	DatoProduktion	By calculation: Year(DatoProduktion)	prodY
S.23	DatoProduktion	By calculation: Month(DatoProduktion)	prodM
S.24	DatoProduktion	By calculation: Day(DatoProduktion)	prodD
S.25	DatoHoldbarhed	By calculation: Year(DatoHoldbarhed)	expiryY
S.26	DatoHoldbarhed	By calculation: Month(DatoHoldbarhed)	expiryM
S.27	DatoHoldbarhed	By calculation: Day(DatoHoldbarhed)	expiryD
S.28	UdtDato	By calculation: Year(UdtDato)	sampY
S.29	UdtDato	By calculation: Month(UdtDato)	sampM
S.30	UdtDato	By calculation: Day(UdtDato)	sampD
S.31	PROJEKTNR		progCode
R.01	PROVENR, COMPONENT (see note)	By calculation: PROVENR paramCode	resultCode
R.02	FAERDIGDATO	By calculation: Year(FAERDIGDATO)	analysisY
R.03	FAERDIGDATO	By calculation: Month(FAERDIGDATO)	analysisM
R.04	FAERDIGDATO	By calculation: Day(FAERDIGDATO)	analysisD
R.06	COMPONENT, MetodeKode (see note)	Mapping table EFSAstof	paramCode
R.07	COMPONENT, MetodeKode	Mapping table EFSAstof	paramText

R.09	METODE			anMethRefCode
R.12	Akkrediteret	(see note)	Mapping table EFSAmdstat	accredProc
R.13	UNITS	(see note)	Mapping table EFSAunit	resUnit
R.18	NUMBER_VALUE			resVal
R.21	SD			resValUncertSD
R.25	UNITS		Mapping table EFSAunit	exprRes
R.27	NUMBER_VALUE, LOQ	(see note)	By calculation in accordance with the Guidance Document	resType
R.28	MRL	(see note)		resLegalLimit
R.30	Konkl1, Konkl2	(see note)		resEvaluation

Notes:

R.01	resultCode	Internally (through EFSAstof), if paramCode= [not in list] then paramCode is maintained with a unique suffix, which is removed before writing to XML. Thus resultCode is unique also for substances not in PARAM catalogue, while XLM file conforms to PARAM catalogue.		
R.06	paramCode	See resultCode; At present two [not in list] codes are defined: RF-XXXX-XXX-XXX-HEX-B (cis-Heptachlorepoxyde) and RF-XXXX-XXX-XXX-HT2-T2 (HT-2 and T-2 toxins (sum))		
R.12	accredProc	Mapped through EFSAstof by (COMPONENT, MetodeKode) in the synthesis of individual result-lines for non-detected residues in certain multi-methods		
R.13	resUnit	Special handling of units based on fat or dry matter		
R.27	resType	At present only dependency of LOQ has implemented. Dependency of LOD, CCA, CCB or resQualValue will be implemented when relevant.		
R.28	resLegalLimit	This information may not be available for all projects. Information will be available for pesticide residues.		
R.30	resEvaluation	Konkl1 and Konkl2 may contain codes for evaluation. Some are relevant for EFSA, some are not. A list of codes, relevant for EFSA, is maintained. If one of these is used, SAS code will ensure that this code is placed in Konkl1. In case two EFSA-relevant codes are found, the result is flagged for validation and decision; only one is transmitted to EFSA.		

Component + Method level (Component and method are defined at result level)

SSD-ID	Source		Mechanism	SAS dataset: EFSA
L.1	COMPONENT, MetodeKode	-	Mapping table EFSAstof	labCode
L.2	COMPONENT, MetodeKode		Mapping table EFSAstof	labAccred
R.08	COMPONENT, MetodeKode, prodCode		Mapping tables EFSAstof and EFSAptype	paramType
R.10	COMPONENT, MetodeKode		Mapping table EFSAstof	anMethCode
R.11	COMPONENT, MetodeKode		Mapping table EFSAstof	anMethText
R.14	COMPONENT, MetodeKode, FAERDIGDATO		Mapping table EFSAstof	resLOD

R.15	COMPONENT, MetodeKode, FAERDIGDATO	(see note)	Mapping table EFSAstof	resLOQ
R.20	COMPONENT, MetodeKode		Mapping table EFSAstof	resValRecCorr
R.29	COMPONENT, MetodeKode	(see note)	Mapping table EFSAstof	resLegalLimitType

Notes:

R.15	resLOQ	LODs and LOQs are revised per analysis year. Multiple entries of (COMPONENT, MetodeKode) may exist in EFSAstof. Each instance has a first and a last date of validity. Relevant instance is joined based on analysis date (FAERDIGDATO). In some projects, resLOQ may be stored on result level (in LGM). If so, the value from LGM will replace resLOQ from EFSAstof.		
R.29	resLegalLimitType	Information on resLegalLimitType cannot be retrieved at result level. Coding at Component/Method level will be correct in most cases. So far, information on resLegalLimit has only been available for pesticide projects.		

Project level (Project is defined at result level)

SSD-ID	Source		Mechanism	SAS dataset: EFSA
S.32	PROJEKTNR		Mapping table EFSAproj	progLegalRef
S.33	PROJEKTNR		Mapping table EFSAproj	progSampStrategy
S.34	PROJEKTNR	(see note)	Mapping table EFSAproj	progType
S.35	PROJEKTNR	(see note)	Mapping table EFSAproj	sampMethod
S.39	PROJEKTNR	(see note)	Mapping table EFSAproj	sampPoint

Notes:

S.34	progType	For pesticide projects, progType from EFSAproj is adjusted by complex mechanism to achieve information for samples belonging to the EU coordinated programme.		
S.35	sampMethod	Information on sampMethod cannot be retrieved at result level. Coding at project level will be correct in many cases (especially where reference can be made to legislation), but there will be exceptions in the domain of some chemical contaminants. Can be entered manually in extra added field in LIMS extract.		
S.39	sampPoint	Information on sampPoint cannot be retrieved at result level. Coding at project level will be correct in many cases, but there will be exceptions in the domain of some chemical contaminants (the information is stored in LIMS at sample level (inspection point register), but DVFFV has decided that lack of resources prevent coding for the retrieval at present. Information in project file can be replaced by manually placed information in LIMS extract.		

Complex mechanism

SSD-ID	Source		Mechanism	SAS dataset: EFSA
S.02	PROVENR, Eksternt prøvenummer		SAS procedure	labSampCode, labSubSampCode

S.34	PROJEKTNR, DKvare, OKO, PROJEKTOPLYSNING2, PROJEKTOPLYSNING3	(see note)	Mapping table EFSAsrc	progType
R.23	(Special procedure)	(see note)		moistPerc
R.24	(Special procedure)	(see note)		fatPerc

Notes:

S.34	progType	Implemented through project file; however, pesticide calls require complex handling.		
R.23	moistPerc	moistPerc is not part of the LIMS sample description. If a method for determination of moistPerc has been used for the sample, the result from this analysis will be collected (the result is identified through EFSAsstof where methods for determination of dry matter are listed; the result is joined to all other results for that sample through SAS code). moistPerc = 100 - dry matter%.		
R.24	fatPerc	fatPerc is not part of the LIMS sample description. If a method for determination of fatPerc has been used for the sample, the result from this analysis will be collected (the result is identified through EFSAsstof where methods for determination of fat are listed; the result is joined to all other results for that sample through SAS code)		

Constants

SSD-ID	Source		Mechanism	SAS dataset: EFSA
S.03	"da"	(see note)	Constant	lang
S.04	"DK"		Constant	sampCountry

Notes:

S.03	lang	Information from free text fields entered into the Danish LIMS will be copied without translation to English. Alternatively English will be used (particularly in fields defined through mapping tables)		
------	------	--	--	--

Not relevant

SSD-ID	Source		Mechanism	SAS dataset: EFSA
S.05	Not used. Information not relevant			sampArea
S.07	Not used. Information not relevant			origArea
S.11	Not used. Information not relevant			procArea
R.05	Not used. Information not relevant			EFSAParamCode

Not implemented

SSD-ID	Source		Mechanism	SAS dataset: EFSA
L.3	Not implemented (see note)			labCountry
O.1	Not implemented (see note)			localOrg
O.2	Not implemented (see note)			localOrgCountry
R.16	Not implemented (see note)			CCalpha
R.17	Not implemented (see note)			CCbeta
R.22	Not implemented (see note)			resValUncert
R.26	Not implemented (see note)			resQualValue

R.31	Not implemented (see note)	Mapping table EFSAeval	actTakenCode
R.32	Not implemented (see note)	Mapping table EFSAeval	resComm

Notes:

L.3	labCountry	When relevant, this field can be implemented through EFSAstof (based on COMPONENT, MetodeKode)
O.1	localOrg	When relevant, this field can be implemented through a constant
O.2	localOrgCountry	When relevant, this field can be implemented through a constant
R.16	CCalpha	When relevant, this field can be implemented by the method used for LOQ
R.17	CCbeta	When relevant, this field can be implemented by the method used for LOQ
R.22	resValUncert	For some methods, the confidence interval for the result is recorded. From this, resValUncert could be derived
R.26	resQualValue	When relevant, this field can be implemented through TEXT_VALUE
R.31	actTakenCode	This information has to be specified by the competent authority. It is not part of the LIMS information. If available, the information can be included by joining on (PRØVENR, COMPONENT). This is done for pesticide calls.
R.32	resComm	This information has to be specified by the competent authority. It is not part of the LIMS information. If available, the information can be included by joining on (PRØVENR, COMPONENT). This is done for pesticide calls.

Not available for automatic mapping

SSD-ID	Source	SAS dataset: EFSA
S.08	Not used. Information not available in general. (see note)	origFishAreaCode
S.09	Not used presently. Information not available. (see note)	origFishAreaText
S.10	Not used. Information not available in general. (see note)	procCountry
S.16	Not used. Information not available in general. (see note)	prodPack
S.18	Not used. Information not available in general. (see note)	prodBrandName
S.19	Not used. Information not available in general. (see note)	prodManuf
S.20	Not used. Information not available in general. (see note)	prodIngred
S.21	Not used. Information not available in general. (see note)	prodCom
S.36	Not used. Information not available in general. (see note)	sampleNum
S.37	Not used. Information not available in general. (see note)	lotSize
S.38	Not used. Information not available in general. (see note)	lotSizeUnit
R.19	Not used. Information not available in general. (see note)	resValRec

Notes:

The mapping described below must at present be done manually, and is quit time consuming.

Special coding in SAS has been used (i.e. pesticides) where strict guidelines for recording in LIMS are established and followed.

S.08	origFishAreaCode	LIMS system is designed to allow for project specific information in up to four free text fields. If specified in project description, one of these fields could be used to record origFishAreaCode as free text.
S.09	origFishAreaText	Where relevant, this information has been copied manually from one of several fields to an added field in the LIMS extract. Sometimes ICES codes ⁸ will be recorded.

⁸ <http://www.ices.dk/indexfla.asp>

S.10	procCountry	LIMS system is designed to allow for project specific information in up to four free text fields. If specified in project description, one of these fields could be used to record procCountry. In some cases, origCountry is used to store information on procCountry in LIMS.
S.16	prodPack	LIMS system is designed to allow for project specific information in up to four free text fields. If specified in project description, one of these fields could be used to record prodPack as free text.
S.18	prodBrandName	LIMS system is designed to allow for project specific information in up to four free text fields. If specified in project description, one of these fields could be used to record prodBrandName as free text.
S.19	prodManuf	LIMS system is designed to allow for project specific information in up to four free text fields. If specified in project description, one of these fields could be used to record this information as free text.
S.20	prodInged	At present, it is unclear how this information is stored in the LIMS, except that when available it will be in the form of free text description in Danish.
S.21	prodCom	LIMS system is designed to allow for project specific information in up to four free text fields. If specified in project description, one of these fields could be used to record prodCom as free text.
S.36	sampleNum	LIMS system is designed to allow for project specific information in up to four free text fields. If specified in project description, one of these fields could be used to record sampleNum as free text.
S.37	lotSize	LIMS system is designed to allow for project specific information in up to four free text fields. If specified in project description, one of these fields could be used to record lotSize as free text.
S.38	lotSizeUnit	LIMS system is designed to allow for project specific information in up to four free text fields. If specified in project description, one of these fields could be used to record lotSizeUnit as free text.
R.19	resValRec	This information is not stored in LIMS